Guest Editorial

ITEA Journal 2010; 31: 161–164

Rigor and Objectivity in T&E

J. Michael Gilmore, Ph.D.
Director, Operational Test and Evaluation,
Office of the Secretary of Defense, Washington, DC

The Director of Operational Test and Evaluation (OT&E) has begun four T&E initiatives. The initiatives are far reaching and will influence the conduct of T&E across the defense establishment. The initiatives begin with support for our forces: "T&E is to help get the capability needed by our fighting forces to them as quickly as possible." The initiatives extend the boundaries for testers to engage in areas not typically the domain of T&E, such as "Review requirements as they are developed to assess whether they are unambiguous, testable, and relevant to accomplishing missions in combat." They promote new methods to improve T&E efficiency and effectiveness, through integrated testing. Finally, they sustain the important past T&E priorities of working with developmental testers to incorporate a reliability growth curve or software failure profile, reliability tests during development, and evaluation of reliability growth and reliability potential during development in the Test and Evaluation Master Plan (TEMP).

In my confirmation hearing before the U.S. Senate last June, I committed to provide objective evaluations of the effectiveness, suitability, and survivability of weapon systems based on realistic operational testing, with my goal to ensure that the men and women in uniform are provided weapons that they can be confident will work. One senator asked how I defined robust testing. I responded that robust testing is the testing needed to provide operators with high confidence that they understand what the system will do and will not do. And of course, acquisition decision makers must have high confidence to enable them to make proper decisions prior to fielding systems. We can achieve both by ensuring *Rigor and Objectivity in Testing and Evaluation (T&E)*.

J. Michael Gilmore, Ph.D.

The Undersecretary of Defense for Acquisition, Technology, and Logistics recently said, of a particular acquisition program, that the cost estimate on which he relies has a 50 percent chance of being wrong. The Weapon System Acquisition Reform Act now requires 80 percent confidence, or the justification for selecting a confidence level of less than 80 percent. In T&E, let us go about our work, so we consistently provide information on which all users can have high confidence.

I believe the T&E community not only has a legislated responsibility for planning and reporting on tests but also has a broader responsibility: to improve the chances that acquisition systems actually meet the needs of soldiers, sailors, airmen, and marines; in other words, "weapons that work when needed."

After being sworn in on September 23, 2009, and consulting with the senior management of Director, Operational Test & Evaluation (DOT&E), I identified four T&E initiatives and announced them in a memorandum dated November 24, 2009. The memorandum is on the DOT&E Web page available at http://www.dote.osd.mil/.

The four initiatives build upon our proven capability to provide rigorous, objective, and clear information in order to

1. Field New Capability Rapidly,
2. Engage Early to Improve Requirements,
3. Integrate Developmental, Live Fire, and Operational Testing, and
4. Substantially Improve Suitability Before Initial Operational Test and Evaluation (IOT&E).

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
JUN 2010		**00-00-2010 to 00-00-2010**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Rigor and Objectivity in T&E	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Office of the Secretary of Defense ,Operational Test and Evaluation,Washington,DC,20301**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**4**	

These initiatives subsume priorities and metrics that have guided the DOT&E organization for the past several years and formed the basis of its annual reports to Congress.

Initiative 1: Field New Capability Rapidly

Secretary of Defense Robert Gates has made clear that his top priority is to get the capability needed by our fighting forces to them as quickly as possible. For T&E to take an initiative supporting the Secretary's top priority should not surprise anyone. The T&E community has been both helpful and unobtrusive in rapidly fielding new capability. We need consider only the Mine Resistant Ambush Protected (MRAP) vehicle as an example of such outstanding T&E support for rapid fielding.

To extend DOT&E's efforts to support rapid fielding as far as possible, the DOT&E staff will review all programs to identify candidates for early fielding or accelerated testing. After many years at war, it appears that there may not be too many systems left to accelerate. But, if testing has already confirmed that a system would be effective and suitable in current theaters of operation, those findings will be identified. If only a small amount of testing remains in order to determine effectiveness and suitability, we will identify opportunities for acceleration of that testing.

As they seek proactively to be involved in early fielding initiatives, the DOT&E staff will assess whether planned testing will be sufficient to identify fully the capabilities and limitations of the program being fielded. The DOT&E staff will also identify opportunities to streamline T&E procedures and processes to support early fielding initiatives. As appropriate, staff will communicate for action their assessments and those opportunities to program offices, the Operational Test Agencies (OTAs), and the DOT&E leadership.

The DOT&E staff should be flexible with respect to T&E procedures to see if they can be expedited. I expect DOT&E staff to be involved in early fielding initiatives, and help—not hinder. If you find we are doing otherwise, please let me know.

The feedback loop from fielding to program development, and later to testing, needs to be strengthened, particularly for rapid fielding initiatives. Thus, the DOT&E staff must work with the OTAs to identify and communicate critical problems with fielded equipment in need of immediate repair to program offices and DOT&E leadership.

The commitment to contribute to the rapid fielding of new capability to our forces will include the T&E of individual systems, efforts to find cases where the delivery of capability can be accelerated, determining that test–fix–test cycles are planned and occur, advocating use of quick reaction testing (QRT) to help develop Tactics, Techniques, and Procedures (TTPs), and the effective and rapid communication of test results to our commanders in the field.

We have seen success in QRT. A recent example is documented in a letter from the Commander, U.S. Central Command, complimenting QRT of procedures for the Joint Entry Control Point/Use of Force Handbook as a model for delivering a timely solution to our warfighters.

T&E must match the commitment of the warfighter. This is our top priority, and it must drive the pace of our daily work.

Initiative 2: Engage Early to Improve Requirements

I probably do not need to tell testers that requirements are seldom if ever perfect. The T&E community has the expertise to help the requirements process. We need to help, and to do so early on in the process—"left" of where testers traditionally engage in the acquisition process. We must do all we can to ensure that systems have realistic, relevant, and testable requirements and to identify programs that fail to meet that standard. To accomplish this initiative, DOT&E will

- Review requirements as they are developed to assess whether they are unambiguous, testable, relevant to accomplishing missions in combat, and operationally and technically realistic;
- Seek opportunities to be involved in reviews of requirements conducted before those requirements are submitted for consideration within the Office of the Secretary of Defense;
- For each project under oversight, review the Test and Evaluation Strategy (TES) and TEMP to ensure they include testing in realistic operational environments initiated during development and continuing through operational testing. This continuum of realistic testing will place increasing stress on subsystem components before final integration into a "full-up" system, thereby identifying problems when they can be fixed cost-effectively;
- Identify operational concerns to program offices at the earliest possible time in order to resolve them in a timely manner;
- Identify test-critical resource shortfalls;
- Ensure that testing in a joint environment is included in TESs and TEMPs where appropriate and feasible;
- Ensure that developers and the operational community share a clear, common understanding

of the planned Concept of Operations (CONOPS) or identify for action by DOT&E leadership inconsistencies in those views. If the CONOPS is not available, work to ensure a representative set of CONOPS is included in TESs and TEMPs; and
- Identify when programs lack a Reliability, Availability, Maintainability-Cost Report providing the rationale for meeting reliability requirements.

We must all strive for requirements that represent mission capability and not system technical specifications—leave those to the system engineers. DOT&E has four action officers who participate in the Joint Chiefs of Staff J8 Functional Capability Boards (part of the joint staff requirements process). I expect them to get the right people engaged to advise the requirements process.

Further, we know the Program Manager does not have control over all the other systems and interfaces needed for his program's mission success. To preclude a limited focus, which short changes the end user, and to step up to the commitment I made to the Senate in my confirmation, I have issued a clarifying policy to make clear that DOT&E will always evaluate a system in the mission context…even when it extends beyond the focus of the Program Manager. Just as important, testers must understand the intended operational context. Only with that information can we properly structure test environments.

Initiative 3: Integrate Developmental, Live Fire, and Operational Testing

Integrated testing is now Department of Defense policy. The legal requirement for a dedicated operational test is also clear. Nonetheless, separation of developmental and operational testing has caused difficulties in the development process that have been documented by a Defense Science Board and the National Academies of Science. Most notably is the lack of operational realism in early testing. Failure modes are hidden and performance limitations become evident only at the end of a program when fixing the problems is expensive, time-consuming, and, often, simply not possible. Therefore, DOT&E action officers will work with their counterparts in the office of Developmental Test and Evaluation (DT&E) and the program offices to incorporate integrated testing into TESs and TEMPs.

The DOT&E staff, as part of its determination of TES and TEMP adequacy, will assess whether the realism included in the early testing is adequate to identify the factors key to understanding whether a new system will actually provide improved military capability, as well as those factors that are not key. Identifying these key factors and screening out unimportant factors is essential to constructing the initial operational test.

In May 2009, the OTA Commanders and DOT&E endorsed Design of Experiments (DOE) as an important means to achieve integrated testing. DOE provides the scientific and rigorous method to plan and execute tests, and evaluate their results. DOE is currently used in some programs to construct individual test events. The DT&E and OT&E offices are working with the OTAs and Developmental Test Centers to apply DOE across the whole developmental and operational test cycle for a program.

Whenever possible, our evaluation of test adequacy must include a rigorous assessment of the confidence level of the test, the power of the test, and the breadth of the test—how well it spans the operational envelope of the system. DOE makes that assessment possible. DOE also will allow DOT&E to make rigorous and objective statements of the confidence we have in the results of the testing. Integrated testing and DOE may save resources, but as the OTA Commanders and DOT&E agreed, it may require more resources to achieve the rigor desired.

I have great expectations for integrated testing and suggest that we all need to work hard to apply it in the planning and execution of our test programs. I have started within DOT&E a task specifically designed to contribute to the whole community's understanding of how to use DOE to do this.

The Office of the Secretary of Defense is helping by collecting best practices for using DOE to enable integrated testing. The OTA Commander's Conference will continue to monitor our progress. In addition, the Defense Acquisition University is developing a continuous learning module to complement their resident courses.

Initiative 4: Substantially Improve Suitability Before IOT&E

Reliability is an essential consideration in our evaluation of system suitability; our forces deserve weapons that work whenever needed. We will encourage programs to have a reliability growth program, and we will track reliability growth during development. DOT&E will

- Assess at appropriate milestones whether programs meet the requirement to have a reliability growth program and identify for action by DOT&E leadership cases where this requirement is not met;

- Work with developmental testers to incorporate a reliability growth curve or software failure profile, reliability tests during development, and evaluation of reliability growth and reliability potential during development in the TEMP; and
- Work with developmental testers to ensure data from the test program are adequate to enable prediction with statistical rigor of reliability growth potential and expected IOT&E results. The rigor should be sufficient to calculate the probabilities of accepting a bad system and rejecting a good system, and those probabilities should be used to plan IOT&E.

For new or restructured programs, DOT&E will not approve TESs and TEMPs lacking a reliability growth curve or software failure profile in order to provide assurance that the system can demonstrate attainment of its reliability requirements.

Summary

We must engage in rigorous and objective testing. We must ensure operators have high confidence in their systems—what they will do and will not do—and arm acquisition decision makers with high confidence to make proper decisions prior to fielding systems. I challenge you to join me to field capability rapidly; engage early to improve requirements; integrate developmental, live fire, and operational testing; and, substantially improve suitability before IOT&E. ❑

Dr. J. Michael Gilmore was sworn in as Director of Operational Test and Evaluation on September 23, 2009. A Presidential appointee confirmed by the United States Senate, he serves as the senior advisor to the Secretary of Defense on operational and live fire test and evaluation of Department of Defense weapon systems. Previously Dr. Gilmore was the assistant director for National Security at the Congressional Budget Office (CBO). Dr. Gilmore is a former Deputy Director of General Purpose Programs within the Office of the Secretary of Defense, Program Analysis and Evaluation (OSD[PA&E]). Dr. Gilmore's service with Program Analysis and Evaluation covered 11 years. Earlier, Dr. Gilmore worked at the Lawrence Livermore National Laboratory; Falcon Associates; and McDonnell Douglas Washington Studies and Analysis Group where he became manager, electronic systems company analysis. Dr. Gilmore is a graduate of Massachusetts Institute of Technology, Cambridge, Massachusetts, where he earned a bachelor of science degree in physics. He subsequently earned master of science and doctor of philosophy degrees in nuclear engineering from the University of Wisconsin, Madison, Wisconsin. E-mail: mike.gilmore@osd.mil